

Методы обучения с подкреплением для решения задач сборки

Юнес Али

Руководитель: Ющенко А. С.

Московский Государственный Технический Университет имени
Н.Э.Баумана



Содержание

Введение

классическое управление робототехники

глубокое обучение робототехники

Постановка задачи

Направления исследований

Обучения с подкреплением

Алгоритмы обучения с подкреплением

План работы

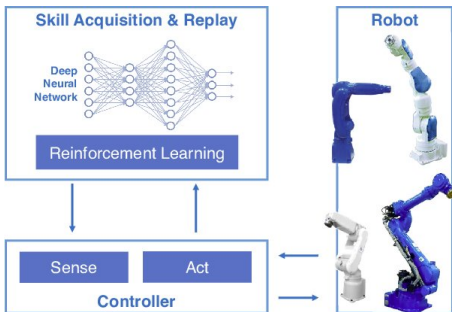
Новизна исследования

Используемые технологии

классическое управление робототехники



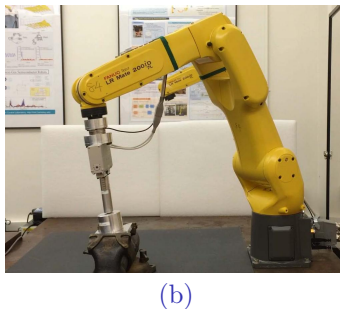
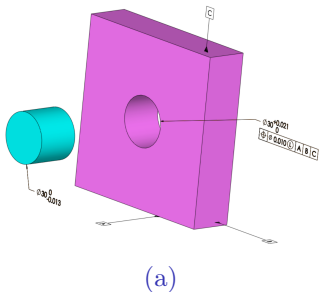
глубокое обучение робототехники



структура системы обучения

Постановка задачи

- ▶ "плотный зазор цилиндрический стержня в отверстии" (Tight clearance cylindrical peg-in-hole task).

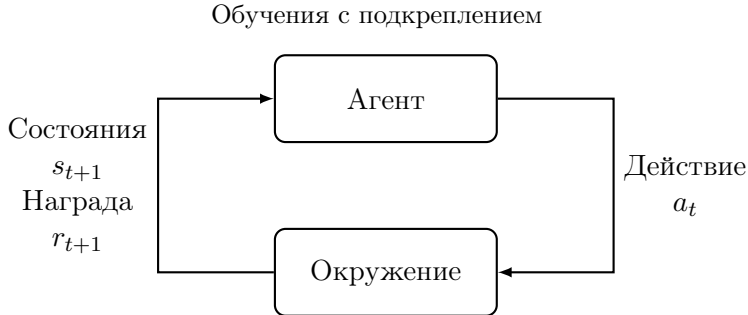


Важность реализации проекта

- ▶ Не зависит от модель робота (с точки зрения алгоритма динамический модель неизвесн)
- ▶ Необходимая Точность для выполнения задачи, превышает точность робота
- ▶ Сверхточные датчики силы и момента → камеру и датчики силы и положения
- ▶ Первый шаг к разработке интеллектуального Робота

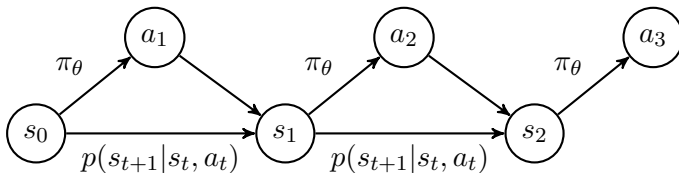
Обучения с подкреплением

- ▶ Является подразделом машинного обучения, изучающим, как агент должен действовать в окружении, чтобы максимизировать некоторый долговременный выигрыш.



Обучения с подкреплением

марковский процесс принятия решений



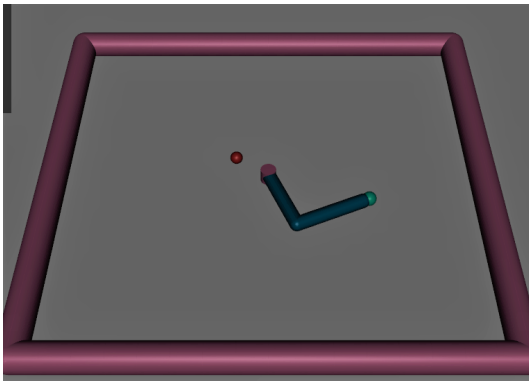
- ▶ $M = S, A, \tau, r$
- ▶ Состояния : $s \in S$
- ▶ Действия : $a \in A$, $a \sim \pi(a_t|s_t)$
- ▶ Вероятность переходов : $p(s_{t+1}|s_t, a_t) \in \tau$
- ▶ Функция вознаграждения : $r : S \times A \rightarrow R$

Алгоритмы обучения с подкреплением

Самые важные современные алгоритмы
(без использования моделей):

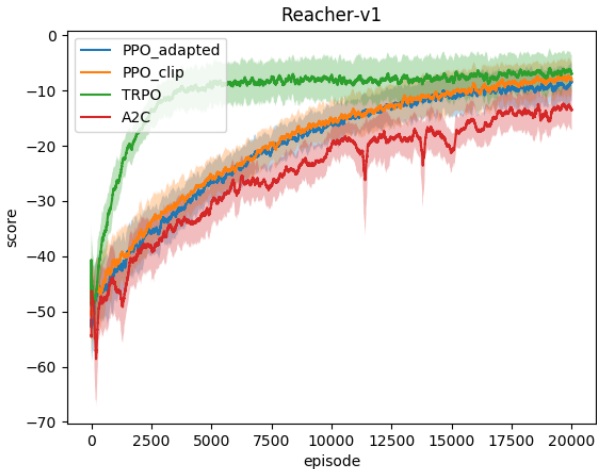
1. Алгоритмы с непрерывными моделями: DQN (2013), DDPG (2015)
2. Алгоритмы градиента стратегии: TRPO (2015), PPO (2017)

Алгоритмы обучения с подкреплением



OpenAI gym environment "Reacher"

Алгоритмы обучения с подкреплением



Новизна исследования

- ▶ В особом случае серийных роботов, мы можем направить робота использовать полезных образцов во время процесса обучения.
- ▶ или мы используем гауссовские процессы, которые могут изучить модель среды из нескольких образцов.
- ▶ это направление исследований называется "Data-efficient reinforcement learning"

Используемые технологии

- ▶ Python, Keras, Tensorflow
- ▶ KUKA LBR IIWA 14-r820
- ▶ ROS (Robotic Operatic System)



ROS



План работы

1. Следовать развития темы исследований (Научные статьи)
2. Написать код алгоритмов (TRPO,PPO,DDPG)
3. Подготовка робота и экспериментальной среды
4. Попробовать алгоритм в симуляторе
5. Попробовать алгоритм на реальном роботе
6. Разработка алгоритма обучения

Конец

Спасибо за внимание!